Metal Artifact Reduction in Pelvic Computed Tomography With Hip Prostheses

Comparison of Virtual Monoenergetic Extrapolations From Dual-Energy Computed Tomography and an Iterative Metal Artifact Reduction Algorithm in a Phantom Study

Kai Higashigaito, MD, Florian Angst, MD, Val M. Runge, MD, Hatem Alkadhi, MD, MPH, EBCR, and Olivio F. Donati, MD

Objective: The aim of this study was to directly compare metal artifact reduction (MAR) of virtual monoenergetic extrapolations (VMEs) from dual-energy computed tomography (CT) with iterative MAR (iMAR) from single energy in pelvic CT with hip prostheses.

Materials and Methods: A human pelvis phantom with unilateral or bilateral metal inserts of different material (steel and titanium) was scanned with third-generation dual-source CT using single (120 kVp) and dual-energy (100/150 kVp) at similar radiation dose (CT dose index, 7.15 mGy). Three image series for each phantom configuration were reconstructed: uncorrected, VME, and iMAR. Two independent, blinded radiologists assessed image quality quantitatively (noise and attenuation) and subjectively (5-point Likert scale). Intraclass correlation coefficients (ICCs) and Cohen κ were calculated to evaluate interreader agreements. Repeated measures analysis of variance and Friedman test were used to compare quantitative and qualitative image quality. Post hoc testing was performed using a corrected (Bonferroni) $P < 0.017$.

Results: Agreements between readers were high for noise (all, ICC ≥ 0.975) and attenuation (all, ICC ≥ 0.986); agreements for qualitative assessment were good to perfect (all, κ ≥ 0.678). Compared with uncorrected images, VME showed significant noise reduction in the phantom with titanium only ($P < 0.017$), and iMAR showed significantly lower noise in all regions and phantom configurations (all, $P < 0.017$). In all phantom configurations, deviations of attenuation were smallest in images reconstructed with iMAR. For VME, there was a tendency toward higher subjective image quality in phantoms with titanium compared with uncorrected images, however, without reaching statistical significance ($P > 0.017$). Subjective image quality was rated significantly higher for images reconstructed with iMAR than for uncorrected images in all phantom configurations (all, $P < 0.017$).

Conclusions: Iterative MAR showed better MAR capabilities than VME in settings with bilateral hip prosthesis or unilateral steel prosthesis. In settings with unilateral hip prosthesis made of titanium, VME and iMAR performed similarly well.

Key Words: algorithms, artifacts, hip prosthesis, metals, phantoms, imaging

(*Invest Radiol* 2015;50: 828–834)

Metallic implants, such as hip prostheses, induce fine streak artifacts as well as broader dark or bright band artifacts, consequently obscuring adjacent pelvic organs, potentially masking pathologies in the pelvis and impairing diagnostic confidence in interpretation of pelvic computed tomography (CT).[1]

Received for publication April 16, 2015; and accepted for publication, after revision, May 20, 2015.

From the Institute of Diagnostic and Interventional Radiology, University Hospital Zurich, University of Zurich, Zurich, Switzerland.

Conflicts of interest and sources of funding: none declared.

Supplemental digital contents are available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.investigativeradiology.com).

Correspondence to: Hatem Alkadhi, MD, MPH, EBCR, Institute of Diagnostic and Interventional Radiology, University Hospital Zurich, University of Zurich, Raemistrasse 100, 8091 Zurich, Switzerland. E-mail: Hatem.Alkadhi@usz.ch.

ISSN: 0020-9996/15/5012–0828

DOI: 10.1097/RLI.0000000000000191

Various metal artifact reduction (MAR) algorithms exist, some based on single-energy techniques and some on dual-energy (DE) techniques. The latter technique[2–5] is characterized by simulating monoenergetic images. On the basis of the attenuation information obtained from the 2 different (low and high peak kilovoltage) energy spectra, the DE data set is decomposed into a linear combination of mass attenuation coefficients of 2 basis materials. Using these 2 data sets, virtual monoenergetic extrapolation (VME) is performed, and monoenergetic images for specific photon energies can be generated.[2] Kiloelectron volt levels for VME images for MAR for various types of metal implants are relatively high and range between 77 and 141 keV.[3–7]

Another recently introduced MAR technique, iterative MAR (iMAR), is based on single-energy CT acquisition.[8,9] Iterative MAR uses 2 previously introduced MAR algorithms, namely, normalized MAR (NMAR) algorithm and image-based frequency-split MAR (FSMAR). First, images are corrected using the NMAR algorithm. Normalized MAR is an inpainting-based MAR method. Metal affected parts of the sinogram are treated as unreliable and are replaced by data generated by normalized interpolation.[10] Second, images are corrected using FSMAR, which combines the high frequencies of the uncorrected original image, with the low frequency of the NMAR corrected image.[11] These 2 steps are repeated several times using the resulting corrected image as input for the next iteration.

Recent studies have compared VME from DE CT with uncorrected images[3,6] or iMAR with uncorrected images,[8,9] each study showing a beneficial effect of the single tested algorithm. To the best of our knowledge, no study has directly compared VME from DE CT and iMAR in the same data set. A direct comparison of these 2 modern MAR techniques is desirable, however, to evaluate which technique performs better in pelvic CT of patients with metal implants, and specifically in this study, hip prostheses.

Thus, the purpose of this study was to directly compare VME from DE CT with iMAR in a phantom study simulating the human pelvis, including 2 types of unilateral and bilateral hip prostheses.

MATERIALS AND METHODS

Phantom

A commercially available phantom (Electron Density Phantom 062M; Computerized Imaging Reference Systems, Norfolk, VA) was used to simulate the human pelvis. The phantom (diameter of 330 × 270 mm) consisted of water and tissue equivalent epoxy resins with cylindrical inserts (diameter of 30 mm) positioned at 17 locations (Fig. 1). The available inserts were arranged to simulate the human pelvis. The parenchymal tissue density insert (location 1) in the center of the phantom was chosen to simulate the prostate and the bone density insert laterally on each side (locations 10 and 14) to simulate the femoral bones. To simulate unilateral hip prosthesis, a metal insert replaced the bone density insert at location 14. To simulate bilateral hip prostheses, both bone density inserts (locations 10 and 14) were replaced by metal inserts. Two different materials, steel and titanium, were used to simulate the

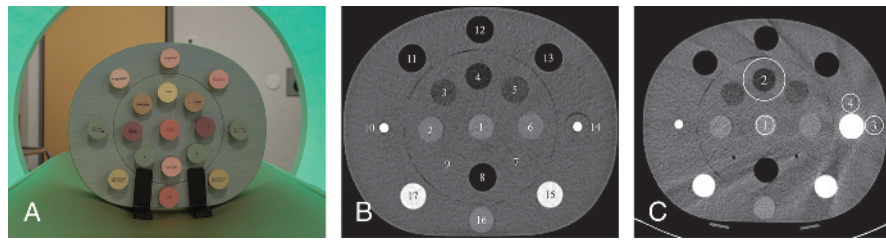

FIGURE 1. Photo of the phantom without metal inserts (reference configuration, Ref) placed on the CT scanner table (A). Computed tomography image of the phantom without metal inserts (reference configuration, Ref) (B). Inserts have the following attenuation: insert 1 and 16, liver equivalent density; 2 and 6, muscle equivalent density; 3 and 5, breast equivalent density; 4, fat equivalent density; 7 and 9, water equivalent density; 8, 11, 12, and 13, lung equivalent density; 10 and 14, solid dense bone equivalent density; 15 and 17, solid trabecular bone equivalent density. Uncorrected CT image of the phantom simulating unilateral titanium hip prosthesis (1Ti) (C). Regions of interest were placed into insert 1 (ROI_1), around insert 4 (ROI_2), lateral (ROI_3), and anterior (ROI_4) to insert 14.

2 main materials used for hip prostheses. Thus, the phantom was scanned with 5 different insert configurations: (1) without any metal inserts to obtain reference images (phantom abbreviation, Ref); (2) steel insert in location 14 (phantom abbreviation, 1Fe); (3) titanium insert in location 14 (abbreviation, 1Ti); (4) steel inserts in locations 10 and 14 (abbreviation, 2Fe); (5) titanium inserts in locations 10 and 14 (abbreviation, 2Ti).

CT Image Acquisition

Scans were performed using a third-generation 192-slice dual-source CT (Somatom Force; Siemens Healthcare, Forchheim, Germany). The phantom was placed into the isocenter of the scanner gantry and was fixed on the table to avoid movement during replacement of the metal inserts. Scans were first performed in single-energy (SE) mode followed by DE imaging. Single-energy data acquisitions were performed with the following parameters: tube voltage, 120 kVp; effective tube current-time product, 108 mA; slice acquisition, 2 × 192 × 0.6 mm; pitch, 1; and rotation time, 1 second. Dual-energy data acquisition was performed with the following parameters: tube A operating at 150 kVp using a tin filter and an effective tube current-time product of 138 mA and tube B operating at 100 kVp and an effective tube current-time product of 69 mA; slice acquisition, 2 × 128 × 0.6 mm; pitch, 0.5; and rotation time, 0.5 second.

All SE and DE scans were performed with a constant radiation dose by adjusting the tube current-time products resulting in a volume CT dose index ($CTDI_{vol}$) of 7.15 mGy. This $CTDI_{vol}$ of 7.15 mGy was determined based on the calculation of the mean $CTDI_{vol}$ of 100 previously acquired, clinically indicated abdominopelvic CT scans of patients performed on the same scanner (54 male, 46 female; 3 unilateral hip prostheses; mean age, 59.2 ± 15.8 years; mean BMI, 25.5 ± 5.3 kg/m^2; mean anteroposterior diameter at the hip, 225 ± 29 mm; mean lateral diameter at the hip, 384 ± 51 mm).

CT Image Reconstruction

The Ref configuration without metal rods was reconstructed using advanced modeled iterative reconstruction (ADMIRE) at a strength level of 3. For all other phantom configurations (1Ti, 2Ti, 1Fe, and 2Fe), 3 image series for each configuration using different reconstruction algorithms were reconstructed: first, uncorrected images without an MAR algorithm with SE data (ADMIRE, strength level 3); second, using the same SE data set with iMAR; and third, using the DE data set with VME images. All reconstructed images were transferred to the picture archiving and communication system (IMPAX 6.5.5; Agfa HealthCare, Mortsel, Belgium) of our hospital.

Uncorrected Images

Using the SE data set, images were reconstructed with ADMIRE (strength level 3) at a slice thickness of 2 mm and increment of 1.6 mm using a medium-smooth soft tissue kernel (Br36).

Iterative MAR

The principles of the iMAR algorithm have been previously published in detail.[8,9] In brief, iMAR uses the combination of NMAR[10] and FSMAR image reconstructions.[11] All iMAR images were reconstructed using a prototype software version (ReconCT, version 13.6.1.0; Siemens) installed on an offline workstation (Xeon CPU dual-core 2.8 GHz, 32 GB RAM; Intel, Santa Clara, CA). This prototype software version offers a selection of presets with optimized reconstruction parameters (such as number of iterations, threshold level for metal segmentation, and filter parameter for frequency-split operations) depending on the used metal implants (eg, hip prosthesis, dental filling, extremity implants). For our study, all iMAR images were reconstructed using the hip prosthesis preset. Images were reconstructed at a slice thickness of 2 mm, increment of 1.6 mm, using a medium-smooth soft tissue kernel (Br36s), field of view of 400 mm, and image matrix of 512 × 512.

Virtual Monoenergetic Extrapolations from Dual-Energy CT

First, images for each tube (150 and 80 kVp) were reconstructed separately with a slice thickness of 0.75 mm, increment of 0.5 mm using a medium-smooth soft tissue kernel (Qr36), and ADMIRE (strength level 3). A commercially available postprocessing software (Syngo.via, version VA30A; Siemens) with dedicated application (Dual-Energy Monoenergetic application, Siemens) was used to generate virtual monoenergetic extrapolations from DE CT data sets. This algorithm decomposes the attenuation information obtained from the DE scan into a linear combination of mass attenuation coefficients of 2 basis materials. Based on this information, virtual monoenergetic extrapolation is performed and monochromatic images for specific kiloelectron volts calculated. Details for this algorithm have been previously published.[2–4,6] First, virtual monoenergetic extrapolations were reconstructed for virtual monoenergetic energies ranging from 40 to 190 keV using the identical slice thickness, increment, and kernels as mentioned previously. Field of view was set to 350 mm with an image matrix of 512 × 512.

One radiologist (with 10 years of experience) who was not involved in subsequent image analyses selected the best monoenergetic kiloelectron volt value showing the least artifacts (streaking artifacts due to beam hardening and photon starvation) for each phantom configuration.

Quantitative Image Analyses

Two independent radiologists (with 4 and 5 years of experience) assessed quantitative image quality by measuring image noise and attenuation values at 4 defined regions in the phantom using a circular region of interest (ROI) (Fig. 1). The first ROI (ROI_1) was placed into the soft tissue density insert (location 1) in the center of the phantom simulating the prostate. The first ROI (ROI area, 491 mm^2) was placed

within the soft tissue insert, carefully avoiding the adjacent epoxy resins of the phantom. To evaluate the potential introduction of new artifacts tangential to high-contrast objects, a second ROI (ROI_2) was placed anterior to ROI_1 around the center of the insert at location 4 (ROI area, 2485 mm^2). The third and fourth ROIs (ROI_3 and ROI_4, respectively; both ROI area, 491 mm^2) were set into the epoxy resins of the phantom in close proximity to the insert at location 14 to evaluate periprosthetic artifacts. Therefore, ROI_3 was set lateral and ROI_4 anterior to insert 14. All measurements were repeated on 3 different representative slices by both radiologists, resulting in a total of 6 measurements for each defined region. Measurements performed in the Ref phantom (without metal rods) were used as the reference values. To quantify the metal artifact–induced deviation of attenuation values, differences between measured attenuation and the corresponding reference values were calculated.

Signal Intensity Profile

Signal intensity profiles were displayed using a custom-written image viewing and analysis software (ViSi version 1.0; Siemens). In each data set, a horizontally oriented line (length, 330 mm) and a second vertically oriented line were drawn through the center of the phantom (length, 270 mm). The horizontal line was drawn through the middle of the bone/metal inserts (locations 10 and 14) and the soft tissue insert in the center (location 1). The vertical line was drawn through the soft tissue insert in the center (location 1), perpendicularly to the horizontal line. All attenuation values along the lines were recorded, and their respective deviation from the reference line, which was drawn through the reference phantom, was calculated (Fig. 2). To calculate the mean deviation of the signal intensity profile, the attenuation value of each point of measurement was compared with the corresponding reference value. The locations of the metal inserts (locations 10 and 14) were excluded from this analysis.

Qualitative Image Analyses

The same 2 radiologists who performed the quantitative image analyses independently assessed subjective image quality in the same regions as for the quantitative image analyses (ROI_1-ROI_4). A fixed window level of 70 Hounsfield units (HU) and fixed window width of 360 HU were used for analyses. For each of the 4 regions, a 5-point Likert scale was used to assess subjective image quality: score 0 = nondiagnostic, major streak artifacts; 1 = marked artifacts, considerably impaired but still diagnostic image quality; 2 = moderate artifacts, adequate image quality; 3 = minor artifacts, good image quality; and 4 = no artifacts, excellent image quality.

Statistical Analysis

For all ROIs, intraclass correlation coefficients (ICCs) were calculated to evaluate the interreader agreement between the 2 readers regarding image noise and attenuation values. Cohen κ was calculated to evaluate interreader agreement regarding assessment of qualitative image quality. Intraclass correlation coefficient less than 0.69 was defined as poor, ICC between 0.70 and 0.79 as fair, ICC between 0.80 and 0.89 as good, and ICC greater than 0.9 as high.[12] Cohen κ value between 0 and 0.20 was defined as slight agreement, κ value between 0.21 and 0.40 as fair, κ value between 0.41 and 0.60 as moderate, κ value

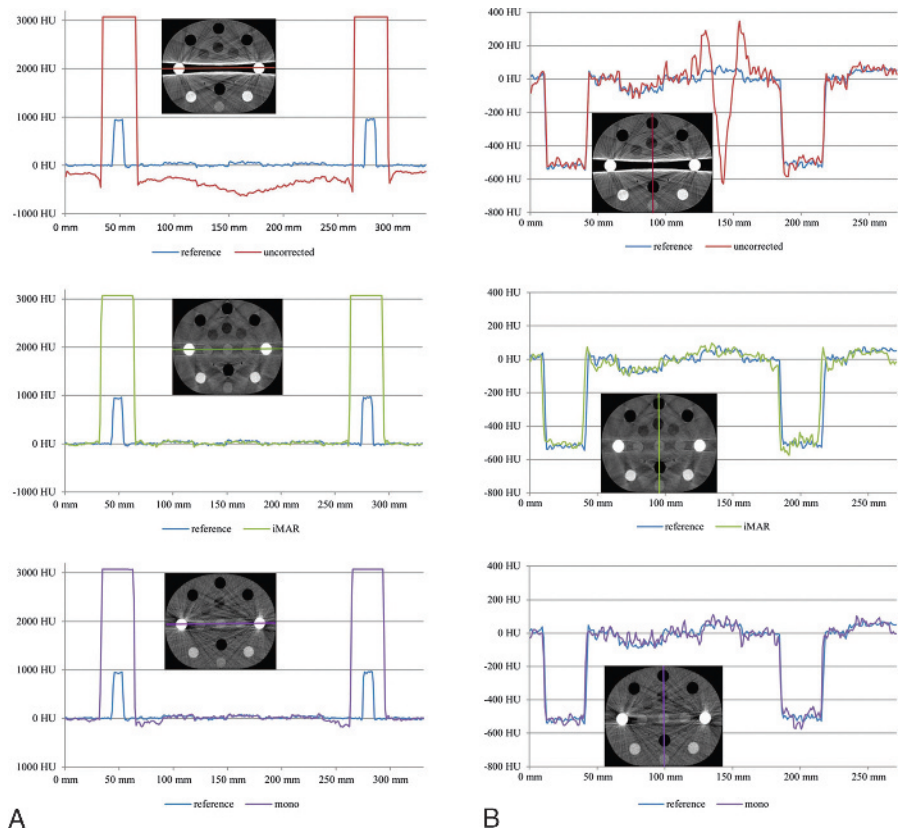

FIGURE 2. Signal intensity profiles from lines drawn horizontally through the center of the 2Ti phantom (A). The profile line crosses the 2 titanium rods resulting in 2 amplitudes at 50 mm and 280 mm. Signal intensity profiles from lines drawn vertically through the center of the 2Ti phantom (B). The profile line crosses the 2 lung density plugs resulting in negative deflection at 20 mm and 200 mm. Note the additional negative deflection of the signal intensity profile line at 150 mm drawn through the uncorrected image representing streak artifacts.

between 0.61 and 0.80 as good, and κ value between 0.81 and 1 as almost perfect agreement.[13]

Normally distributed continuous variables were provided as mean (standard deviation), abnormally distributed variables as median and range. All statistical analyses were performed separately for each phantom configuration, comparing uncorrected images with VME and with images reconstructed with iMAR.

The repeated measures analysis of variance was used to compare image noise between the 3 data sets of each phantom configuration. The paired *t* test was used for pairwise comparison between each reconstruction data set.

The metal artifact–induced deviations of attenuation values from the corresponding reference values of each reconstruction algorithm (uncorrected, VME, iMAR) were compared using the analysis of variance test.

Friedman test was used to compare subjective image quality among the 3 different reconstruction algorithms, and the Wilcoxon signed rank test was used for pairwise comparison between each algorithm.

A 2-tailed *P* value below 0.05 was considered to indicate statistical significance. In case of multiple testing, the *P* value was adjusted according to Bonferroni to a value below 0.017. Statistical analyses were performed using commercially available software (SPSS statistics, version 22, release 22.0.0.1; IBM, Chicago, IL).

RESULTS

Monoenergetic Kiloelectron Volt Values

The following kiloelectron volt values were chosen as those providing the least artifacts: phantom configuration 1Fe, 94 keV; 2Fe, 83 keV; 1Ti, 107 keV; 2Ti, 143 keV (Supplementary Figure 1, Supplemental Digital Content 1, http://links.lww.com/RLI/A216).

Interreader Agreement

Interreader agreement between both readers was high for noise (ICC = 0.997) and attenuation (ICC = 0.996). Interreader agreement for qualitative assessment of image quality was almost perfect (κ = 0.85). Thus, for further analyses, the mean values of the 2 readers were calculated for each ROI.

Quantitative Image Analyses

Image Noise

Significant differences in image noise were found in all phantom configurations and ROIs (all, $P < 0.05$). Differences of image noise and corresponding *P* values between each reconstruction algorithm for each phantom configuration and ROI are listed in Table 1.

A tendency toward decreased image noise as compared with the uncorrected images was found in all phantom configurations using VME, but only in the 1Ti phantom, significance was achieved in all 4 ROIs (all, $P < 0.017$). Mean noise reduction was 23% in 1Ti, 21% in 2Ti, 19% in 1Fe, and 11% in 2Fe (Table 1).

Using iMAR, a significant decrease in image noise was found in all ROIs and phantom configuration as compared with uncorrected images (all, $P < 0.017$) with a mean noise reduction of 23% in 1Ti, 49% in 2Ti, 52% in 1Fe, and 74% in 2Fe (Table 1).

No significant difference in image noise was observed between VME and iMAR in the 1Ti phantom. In all other phantoms, iMAR showed a significant decrease in image noise in all ROIs as compared with VME (all, $P < 0.011$) with a mean noise reduction of 35% in 2Ti, 42% in 1Fe, and 71% 2Fe (Table 1).

Attenuation Values

Mean attenuation values of each ROI and each reconstruction algorithm are listed in Table 2. A significant difference was found among all phantoms and ROIs ($P < 0.05$). Deviations of attenuation are illustrated in Figure 3.

In phantoms with a unilateral hip prosthesis, results for titanium and steel were similar (1Ti and 1Fe). Deviations of attenuation for central ROIs (ROI_1 and ROI_2) were higher using VME than using iMAR (Table 2). Deviations in periprosthetic regions (ROI_3 and ROI_4) were higher for VME images than for iMAR images (Table 2).

TABLE 1. Image Noise in All ROIs and in All Phantoms

		Uncorrected Versus VME		Uncorrected Versus iMAR		VME Versus iMAR*	
Phantom	**ROI**	**Difference of Noise**	***P***	**Difference of Noise**	***P***	**Difference of Noise**	***P***
1Ti	ROI_1	−14%	0.009	−16%	<0.001	−3%	0.233
1Ti	ROI_2	−22%	<0.001	−10%	<0.001	+15%	<0.001
1Ti	ROI_3	−32%	0.006	−28%	0.009	+6%	0.410
1Ti	ROI_4	−26%	0.001	−35%	<0.001	−13%	0.117
2Ti	ROI_1	−88%	<0.001	−92%	<0.001	−39%	<0.001
2Ti	ROI_2	−12%	0.008	−22%	<0.001	−11%	<0.001
2Ti	ROI_3	+3%	0.622	−42%	<0.001	−43%	<0.001
2Ti	ROI_4	+15%	0.015	−39%	<0.001	−47%	<0.001
1Fe	ROI_1	−6%	0.310	−25%	0.003	−20%	0.011
1Fe	ROI_2	−16%	<0.001	−29%	<0.001	−15%	<0.001
1Fe	ROI_3	−16%	0.008	−78%	<0.001	−74%	<0.001
1Fe	ROI_4	−37%	<0.001	−74%	<0.001	−59%	<0.001
2Fe	ROI_1	−2%	0.290	−94%	<0.001	−94%	<0.001
2Fe	ROI_2	+4%	0.034	−46%	<0.001	−47%	<0.001
2Fe	ROI_3	−13%	0.117	−78%	<0.001	−75%	<0.001
2Fe	ROI_4	−34%	<0.001	−78%	<0.001	−67%	<0.001

$P < 0.017$ was considered statistically significant.

*Negative values represent higher noise in VME as compared with iMAR.

ROI indicates region of interest; VME, virtual monoenergetic extrapolation; iMAR, iterative metal artifact reduction.

TABLE 2. Attenuation Value and Standard Deviation in All ROIs and All Phantoms

ROI	Phantom	Mean HU Reference	Mean HU Uncorrected	Mean HU VME	Mean HU iMAR
ROI_1	1Ti	55 (0)	50 (1)	52 (1)	49 (3)
ROI_2	1Ti	−20 (2)	−17 (4)	−12 (2)	−19 (4)
ROI_3	1Ti	2 (2)	−30 (9)	−11 (7)	4 (3)
ROI_4	1Ti	5 (3)	6 (3)	24 (4)	9 (3)
ROI_1	2Ti	55 (0)	−236 (45)	66 (3)	43 (2)
ROI_2	2Ti	−20 (2)	−8 (3)	−10 (2)	−22 (2)
ROI_3	2Ti	2 (2)	−157 (10)	−14 (8)	−3 (3)
ROI_4	2Ti	5 (3)	30 (9)	44 (7)	15 (2)
ROI_1	1FE	55 (0)	29 (2)	43 (3)	49 (2)
ROI_2	1FE	−20 (2)	−16 (4)	−12 (1)	−18 (1)
ROI_3	1FE	2 (2)	−206 (24)	−125 (22)	0 (2)
ROI_4	1FE	5 (3)	108 (18)	45 (3)	11 (3)
ROI_1	2FE	55 (0)	−415 (4)	−368 (13)	42 (2)
ROI_2	2FE	−20 (2)	1 (3)	−1 (1)	−22 (3)
ROI_3	2FE	2 (2)	−369 (23)	−295 (16)	−6 (4)
ROI_4	2FE	5 (3)	152 (14)	67 (8)	14 (4)

ROI indicates region of interest; HU, Hounsfield unit; VME, virtual monoenergetic extrapolation; iMAR, iterative metal artifact reduction.

Phantoms with bilateral hip prosthesis (2Ti and 2Fe) showed similar results for uncorrected images and VME. A different pattern was only observed in ROI_1 of the 2Ti phantom reconstructed with VME, which showed a small deviation from the corresponding reference value (−10 HU). All other measurements of ROI_1 (2Ti and 2Fe of uncorrected images and 2Fe of VME) showed high deviations ranging between 292 to 470 HU (Table 2).

Deviations for ROI 2 were small and ranged between −21 to −12 HU for uncorrected and −10 to 19 HU for VME. The periprosthetic region (ROI_3 and ROI_4) showed high deviations ranging between −148 to 371 HU for uncorrected images and −62 to 297 HU for VME. In contrast, iMAR showed small deviations for all ROIs (Table 2).

TABLE 3. Mean Deviation of the Signal Intensity Profile Compared With the Reference

	1Ti	2Ti	1Fe	2Fe
Uncorrected (horizontal)	27 HU	354 HU	112 HU	564 HU
VME (horizontal)	21 HU	37 HU	57 HU	487 HU
iMAR (horizontal)	16 HU	21 HU	19 HU	26 HU
Uncorrected (vertical)	20 HU	61 HU	31 HU	98 HU
VME (vertical)	20 HU	26 HU	29 HU	99 HU
iMAR (vertical)	20 HU	25 HU	22 HU	30 HU

HU indicates Hounsfield unit; VME, virtual monoenergetic extrapolation; iMAR, iterative metal artifact reduction.

Signal Intensity Profile

Mean deviations of each horizontal and vertical signal intensity profile from the corresponding reference signal intensity profile for each phantom are given in Table 3. An example of signal intensity profiles drawn through the 2Ti phantom is illustrated in Figure 2.

In all phantom configurations, the deviation of the horizontal signal intensity lines was smallest on the images reconstructed with iMAR. Vertical signal intensity lines showed similar deviations for the T1 phantom for all reconstruction algorithms, whereas in all other phantom configurations, deviation was smallest with iMAR.

Qualitative Image Analysis

Image examples from all phantom configurations and reconstruction algorithms are illustrated in Figure 4. All regions of the reference image were rated as having excellent image quality without artifacts (score 4). Scores for each reconstruction algorithm and phantom configuration are given as median and range in Table 4. Subjective image quality was significantly different within all 3-reconstruction

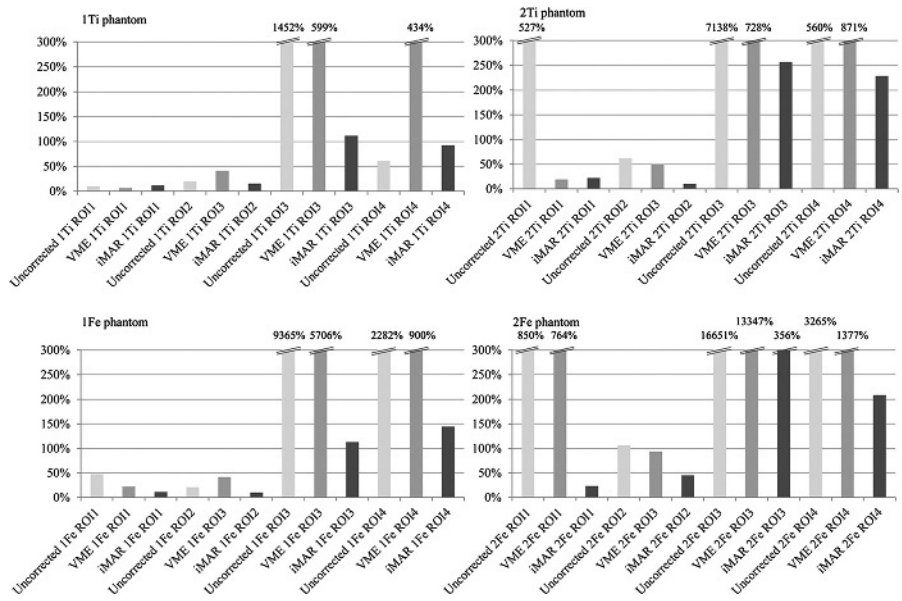

FIGURE 3. Bar graphs demonstrating the relative deviations of attenuation from the reference value (taken from the Ref phantom) for each ROI, phantom configuration, and reconstruction algorithm.

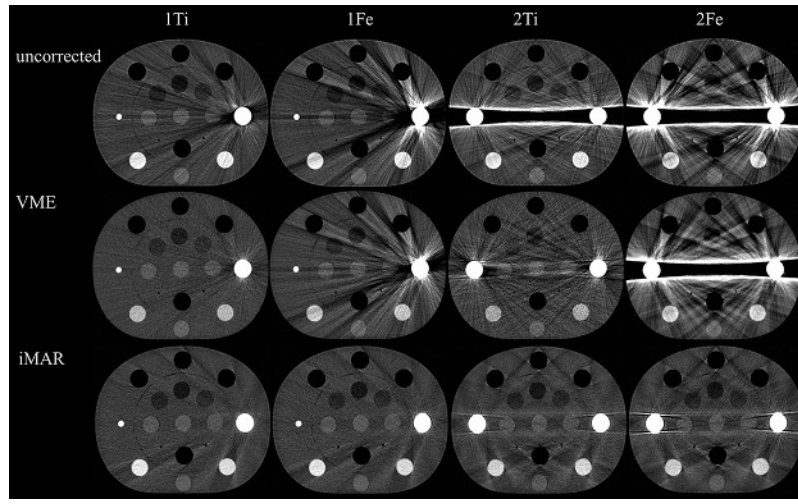

FIGURE 4. Image examples from all phantom configurations and reconstruction algorithms. First column, 1Ti phantom; second column, 1Fe phantom; third column, 2Ti phantom; fourth column, 2Fe phantom. Top panel, Uncorrected images; middle panel, VME; bottom panel, images reconstructed with iMAR.

algorithms for all 4 phantom configurations (1Ti, $P = 0.004$; 2Ti, $P = 0.002$; 1Fe, $P = 0.001$; 2Fe, $P = 0.001$).

Regarding phantom configurations with 1 titanium rod, VME showed a tendency toward higher image quality (median score, 1Ti = 3) as compared with uncorrected images of the corresponding phantom configuration (median score, 1Ti = 2.5); however, statistical significance was not achieved (1Ti, $P = 0.157$). For all other phantom configurations, there was no tendency toward improved image quality with VME as compared with uncorrected images (Table 4). Iterative MAR showed a significant improvement of subjective image quality in all phantom configurations (1Ti, $P = 0.008$; 2Ti, $P = 0.010$; 1Fe, $P = 0.009$; 2Fe, $P = 0.010$) compared with the uncorrected images (Table 4). Image quality using iMAR was significantly higher than using VME in all phantom configurations combinations except in 1Ti.

DISCUSSION

Our study results indicate that iMAR performed better than VME for reducing metal artifacts in CT when 2 hip prostheses were present and in the case of a unilateral prosthesis made from steel. In a configuration with a unilateral titanium prosthesis, both reconstruction algorithms performed comparably well.

Several studies regarding MAR using VME on dual-source or single-source CT have been published. Meinel et al[6] assessed MAR capabilities of VME in 22 patients with various kinds of metal implants, including 5 unilateral and 2 bilateral hip prostheses using energy levels ranging between 100 and 130 keV.[3] As compared with uncorrected images, a significant reduction of artifacts adjacent to the metallic implants was demonstrated by VME. Similar results with VME (105 keV) were demonstrated by Bamberg et al,[1] including 12 patients with unilateral hip prostheses, and Lee et al[14] evaluated the potential of MAR in a phantom study with titanium and stainless steel plates using VME on a single-source CT system using 80 keV for titanium plates and 110 keV for stainless steel plates. Iterative MAR was recently assessed in 41 patients with 32 unilateral and 9 bilateral hip prostheses, focusing on the image quality of the pelvic organs.[6] An increase in objective and subjective image quality was demonstrated for iMAR data sets when compared with uncorrected images. Attenuation, in terms of deviation of CT numbers from reference values in images reconstructed by iMAR, was recently analyzed in a pelvic phantom study with steel and titanium inserts.[5] In images reconstructed with iMAR, CT numbers were shown to deviate significantly less form the reference values as compared with uncorrected images.

The aforementioned studies showed promising results regarding reduction of metal artifacts using either VME or iMAR for hip prostheses. However, no study systematically compared the performance of the 2 MAR algorithms for different scenarios (unilateral vs bilateral hip prosthesis) and for different types of metal (steel and titanium). Different quantitative criteria were used in our study to evaluate the performance of these 2 MAR algorithms. First, in contrast to prior studies, ROIs were placed in different areas of the phantom to evaluate central pelvic (ROI_1 and ROI_2) as well as periprosthetic areas (ROI_3 and

TABLE 4. Subjective Image Quality

				Uncorrected Versus VME	Uncorrected Versus iMAR	iMAR Versus VME
Phantom	**Uncorrected**	**VME**	**iMAR**	***P***	***P***	***P***
1Ti	2.5 (2–3)	3 (2–3)	3 (3–4)	0.157	0.008	0.025
2Ti	1 (0–2)	1 (1–1)	3 (2–3)	1	0.010	0.007
1Fe	1.5 (1–3)	1 (0–2)	3 (3–4)	0.059	0.009	0.010
2Fe	0.5 (0–1)	0 (0–1)	3 (2–3)	0.157	0.010	0.010

Scores are provided as median and range. $P < 0.017$ was considered statistically significant.

VME indicates virtual monoenergetic extrapolation; iMAR, iterative metal artifact reduction.

ROI_4). Second, image noise and deviation of signal intensity lines were used to evaluate the extent of streaking artifacts. Third, deviations of attenuation values compared with the reference values were used to evaluate the HU accuracy of the different MAR algorithms. Using these measures, we found significant reduction of noise in most regions in phantoms with unilateral prostheses from both steel and titanium and for bilateral titanium prostheses using VME from DE CT. Especially in the 1Ti phantom, VME achieved similar results as compared with iMAR without significant difference in qualitative image quality or noise reduction (in 3 of 4 ROIs). In the 2Ti phantom, VME achieved a significant decrease in artifacts only in the ROIs placed in the center of the phantom (ROI_1 and ROI_2) but no significant decrease in artifacts in the periprosthetic regions (ROI_3 and ROI_4). In the 1Fe phantom, VME achieved a minimal but significant decrease in noise in the periprosthetic area (ROI_3 and ROI_4); however, in subjective analysis, there was no improvement in image quality. Iterative MAR showed significant improvement of qualitative and quantitative image quality for all evaluated regions compared with uncorrected images in all phantom configurations. Thus, it can be summarized that iMAR showed better artifact reduction capabilities than VME for the 2Ti, 1Fe, and 2Fe phantoms.

Our results suggest the use of iMAR in all settings with bilateral hip prosthesis regardless of the material used (titanium or steel) and in settings with unilateral hip prosthesis made of steel. In settings with a unilateral hip prosthesis made of titanium, VME and iMAR performed similarly. However, in clinical routine, both the presence and the used material of hip prostheses are frequently unknown before the scan (eg, in trauma patients with suspected pelvic or hip fracture). Considering that iMAR may be applied to SE CT data sets, MAR may be performed retrospectively (ie, after scanning) in routinely acquired scans if considered necessary.

We acknowledge the following study limitations. First, the study was performed using a phantom simulating the human pelvis, thus our results may not be directly applicable in patients. Second, metal inserts simulating the hip prosthesis were oriented along the *z* axis, and a changed spatial orientation of the metal inserts may influence the results of each of the tested MAR algorithms. In addition, we used metal inserts provided with the used phantoms instead of real hip prostheses. The inserts are not tapered, and their diameters may differ from commercially available hip prostheses depending on the location of measurement. Third, our chosen kiloelectron volt levels for VME were lower as compared with prior studies.[3,4,6] However, these studies evaluated the MAR capabilities in the periprosthetic region but not in remote areas of the prosthesis such as in the center of the pelvis. In our study, we evaluated the periprosthetic areas (ROI_3 and ROI_4) as well as areas in the center (ROI_1 and ROI_2) of the phantom. Higher kiloelectron volt levels resulted in visually better MAR in the periprosthetic area; however, new artifacts were introduced in the center of the pelvic phantom using higher kiloelectron volt. Finally, whereas DE CT nowadays can be performed by a number of different single-source and dual-source systems,[15–17] the iMAR algorithm tested herein is limited to a single vendor. In addition, we had prototype software available for performing the iMAR reconstructions from the raw data. However, iMAR was recently released for commercial use including Food and Drug Administration (510K) approval, which implies a growing availability of the technique for various CT scanner generations and types.

In conclusion, iMAR showed better MAR capabilities than VME in settings with bilateral hip prostheses or a unilateral steel prosthesis. In settings with a unilateral hip prosthesis made of titanium, VME and iMAR performed similarly well.

ACKNOWLEDGMENT

The authors thank Karl Ferdinand Loewenich for providing the phantom for this study.

REFERENCES

1. Roth TD, Maertz NA, Parr JA, et al. CT of the hip prosthesis: appearance of components, fixation, and complications. *Radiographics*. 2012;32:1089–1107.
2. Yu L, Leng S, McCollough CH. Dual-energy CT-based monochromatic imaging. *AJR Am J Roentgenol*. 2012;199(suppl 5):S9–S15.
3. Bamberg F, Dierks A, Nikolaou K, et al. Metal artifact reduction by dual energy computed tomography using monoenergetic extrapolation. *Eur Radiol*. 2011;21:1424–1429.
4. Guggenberger R, Winklhofer S, Osterhoff G, et al. Metallic artefact reduction with monoenergetic dual-energy CT: systematic ex vivo evaluation of posterior spinal fusion implants from various vendors and different spine levels. *Eur Radiol*. 2012;22:2357–2364.
5. Han SC, Chung YE, Lee YH, et al. Metal artifact reduction software used with abdominopelvic dual-energy CT of patients with metal hip prostheses: assessment of image quality and clinical feasibility. *AJR Am J Roentgenol*. 2014;203:788–795.
6. Meinel FG, Bischoff B, Zhang Q, et al. Metal artifact reduction by dual-energy computed tomography using energetic extrapolation: a systematically optimized protocol. *Invest Radiol*. 2012;47:406–414.
7. Mangold S, Gatidis S, Luz O, et al. Single-source dual-energy computed tomography: use of monoenergetic extrapolation for a reduction of metal artifacts. *Invest Radiol*. 2014;49:788–793.
8. Axente M, Paidi A, Von Eyben R, et al. Clinical evaluation of the iterative metal artifact reduction algorithm for CT simulation in radiotherapy. *Med Phys*. 2015;42:1170–1183.
9. Morsbach F, Bickelhaupt S, Wanner GA, et al. Reduction of metal artifacts from hip prostheses on CT images of the pelvis: value of iterative reconstructions. *Radiology*.2013;268:237–244.
10. Meyer E, Raupach R, Lell M, et al. Normalized metal artifact reduction (NMAR) in computed tomography. *Med Phys*. 2010;37:5482–5493.
11. Meyer E, Raupach R, Lell M, et al. Frequency split metal artifact reduction (FSMAR) in computed tomography. *Med Phys*. 2012;39:1904–1916.
12. Meyers CR, Blesh TE. *Measurement in Physical Education*. New York, NY: Ronald Press; 1962.
13. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159–174.
14. Lee YH, Park KK, Song HT, et al. Metal artefact reduction in gemstone spectral imaging dual-energy CT with and without metal artefact reduction software. *Eur Radiol*. 2012;22:1331–1340.
15. Morsbach F, Wurnig MC, Müller D, et al. Feasibility of single-source dual-energy computed tomography for urinary stone characterization and value of iterative reconstructions. *Invest Radiol*. 2014;49:125–130.
16. Diekhoff T, Kiefer T, Stroux A, et al. Detection and characterization of crystal suspensions using single-source dual-energy computed tomography: a phantom model of crystal arthropathies. *Invest Radiol*. 2015;50:255–260.
17. Yamada Y, Jinzaki M, Tanami Y, et al. Virtual monochromatic spectral imaging for the evaluation of hypovascular hepatic metastases: the optimal monochromatic level with fast kilovoltage switching dual-energy computed tomography. *Invest Radiol*. 2012;47:292–298.